Asymptotics in Deconvolution Models

Donauer, S.

2009

document version
Publisher's PDF, also known as Version of record

Three

Consistency Results

In this chapter, various consistency results for $\hat{F}_n$ as well as $\hat{h}_n$ are derived. In particular, Hellinger consistency of $\hat{h}_n$ is proven, using techniques from empirical process theory. Uniform strong consistency of $\hat{F}_n$ is deduced from that (Section 3.1). We also discuss the asymptotic behavior of the last point of jump $\hat{S}_n$ of $\hat{F}_n$ (Section 3.2). It turns out that $\hat{S}_n$ stays away from the upper support point of h_0 with probability tending to one under a local condition on g near zero.

3.1 CONSISTENCY OF $\hat{F}_n$ AND $\hat{h}_n$

We consider the same model and use the same notation as in the previous chapter. In particular $F_0 \in \mathcal{F}_{[0,\infty)}$ and $g : [0,\infty) \to [0,\infty)$ as introduced in (2.1.3). With $\|\cdot\|_{L_p}, p \geq 1$, we denote the L_p-norm with respect to Lebesgue measure. Consistency results will be derived with respect to the L_2-, the uniform- and the Hellinger metric where the latter is defined as follows.

DEFINITION 3.1.1 (HELLINGER METRIC).
Let p_1 and p_2 be densities with respect to Lebesgue measure. Then

$$d_{\mathrm{H}}(p_1, p_2) \;=\; \left[\frac{1}{2}\int \left(\sqrt{p_1(x)} - \sqrt{p_2(x)}\right)^2 dx\right]^{1/2}$$

is called the Hellinger distance between p_1 and p_2.

The Hellinger metric is related to the L_1-metric by

$$\|p_1 - p_2\|_{L_1} \;\leq\; 2\sqrt{2}\, d_H(p_1, p_2) \tag{3.1.1}$$

since

$$\begin{aligned}
&\left[\int |p_1(z) - p_2(z)|\, dz\right]^2 \\
&\leq \int \left(\sqrt{p_1(z)} - \sqrt{p_2(z)}\right)^2 dz \cdot \int \left(\sqrt{p_1(z)} + \sqrt{p_2(z)}\right)^2 dz \\
&= 4d_H^2(p_1, p_2)\left(1 + \int \sqrt{p_1(z)p_2(z)}\, dz\right) \;\leq\; 8d_H^2(p_1, p_2)
\end{aligned}$$

using $\int \sqrt{p_1(z)p_2(z)}\, dz \leq 1$. Moreover, if p_1 and p_2 are bounded by a constant $c > 0$,

$$\|p_1 - p_2\|_{L_2} \;\leq\; 2\sqrt{2c}\, d_H(p_1, p_2) \tag{3.1.2}$$

due to

$$\begin{aligned}
\|p_1 - p_2\|_{L_2}^2 &= \int \left(\sqrt{p_1(z)} - \sqrt{p_2(z)}\right)^2 \left(\sqrt{p_1(z)} + \sqrt{p_2(z)}\right)^2 dz \\
&\leq 8cd_H^2(p_1, p_2).
\end{aligned}$$

LEMMA 3.1.2 (HELLINGER CONSISTENCY OF $\hat{h}_n$).
The sequence $\{\hat{h}_n\}_{n=1}^{\infty}$ is consistent with respect to the Hellinger metric, i.e.

$$d_H(\hat{h}_n, h_0) \;\to\; 0 \text{ a.s., as } n \to \infty.$$

Before stating the proof of Lemma 3.1.2, we introduce a few function classes which will appear in the proof:

$$\begin{aligned}\mathcal{H}_1 &:= \{h_1 \mid h_1(z) = g(0)F(z),\, z \geq 0,\, F \in \mathcal{F}_{[0,\infty)}\} \text{ and}\\ \mathcal{H}_2 &:= \left\{h_2 \mid h_2(z) = -\int_0^z F(z-y)g'(y)\,dy,\, z \geq 0,\, F \in \mathcal{F}_{[0,\infty)}\right\}.\end{aligned}$$

Note that $\mathcal{H} = \{g * dF : F \in \mathcal{F}_{[0,\infty)}\}$ is a subset of $\mathcal{H}_1 \ominus \mathcal{H}_2 = \{h_1 - h_2 : h_1 \in \mathcal{H}_1, h_2 \in \mathcal{H}_2\}$ and that every element of $\mathcal{H}_1$ as well as of $\mathcal{H}_2$ is nonnegative, bounded by $g(0)$ and nondecreasing, implying that $\mathcal{H}$ is of uniform bounded variation. Also define the set

$$\mathcal{H}^q := \left\{q \mid q(z) = \frac{h(z)}{h(z) + h_0(z)},\, z \geq 0,\; h \in \mathcal{H}\right\} \tag{3.1.3}$$

whose elements are uniformly bounded by 1.

Proof of Lemma 3.1.2.
Since $\mathcal{H}$ is convex, Lemma 4.5 in van de Geer (2000) can be used to relate the Hellinger metric of $\hat{h}_n$ and h_0 to an empirical process by

$$d_H^2(\hat{h}_n, h_0) \leq 2\int \frac{\hat{h}_n(z)}{\hat{h}_n(z) + h_0(z)}\, d(H_n - H_0)(z),$$

a consequnce of $\Psi_n(\hat{F}_n) \geq \Psi_n((\hat{F}_n + F_0)/2)$. Thus, Hellinger consistency of $\hat{h}_n$ follows if $\mathcal{H}^q$ is an H_0-Glivenko Cantelli class. This will be shown in what follows by first considering $\mathcal{H}_1$ and $\mathcal{H}_2$, then $\mathcal{H}$ and eventually $\mathcal{H}^q$.

From van der Vaart and Wellner (2000, p.125) we know that a function class $\mathcal{A}$ is Glivenko-Cantelli if there exists an integrable envelope and if

$$H(\delta, \mathcal{A}, \|\cdot\|_{L_1(H_n)}) = o_p(n) \text{ for all } \delta > 0, \tag{3.1.4}$$

where H denotes the entropy of $\mathcal{A}$ with respect to the $L_1(H_n)$-norm (see Definition B.1). Note that the function identically equal to $g(0)$ serves as an integrable envelope for the class $\mathcal{H}_1$. We also have

$$\begin{aligned}H\left(\delta, \mathcal{H}_1, \|\cdot\|_{L_1(H_n)}\right) &= H\left(\frac{\delta}{g(0)}, \mathcal{F}_{[0,\infty)}, \|\cdot\|_{L_1(H_n)}\right)\\ \leq\; H\left(\frac{\delta}{g(0)}, \mathcal{F}_{[0,\infty)}, \|\cdot\|_{L_\infty}\right) &\leq c\,\delta \log n = o_p(n)\end{aligned}$$

for some $c > 0$ and all $\delta > 0$ and where L_∞ denotes the supremum norm on $[0,\infty)$. Here the first inequality comes from Lemma 2.1 in van de Geer (2000), which relates covering numbers with respect to different norms. The entropy of a class of monotone functions with respect to some discrete measure is stated in Lemma 2.2 in van de Geer (1993) and yields the last inequality.

For $\mathcal{H}_2$ we can derive (3.1.4) in an analog way, replacing $\mathcal{F}_{[0,\infty)}$ in the last display by $\mathcal{H}_2/g(0) = \{h_g \,|\, h_g(z) = h(z)/g(0),\ z \geq 0,\ h \in \mathcal{H}_2\}$ since also the elements of $\mathcal{H}_2$ are nondecreasing and bounded by $g(0)$.

Define the continuous function $\varphi_1 : \mathbb{R}^2 \to \mathbb{R}$, $\varphi_1(x,y) = x - y$ and apply Theorem 3 from van der Vaart and Wellner (2000), a preservation theorem for Glivenko-Cantelli classes. For that, interpret x and y as the sets $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, and conclude that $\mathcal{H} \subset \varphi_1(\mathcal{H}_1, \mathcal{H}_2)$ is a Glivenko-Cantelli class, where again the constant function $g(0)$ serves as an integrable envelope function.

To finish the proof, we deduce the Glivenko-Cantelli property for $\mathcal{H}^q$ from the corresponding property of $\mathcal{H}$. The following decomposition, valid for any $\varepsilon > 0$, helps to overcome the problem that h_0 causes in the denominator once it becomes really small:

$$
\begin{aligned}
\sup_{\mathcal{H}^q} \left| \int_{[0,\infty)} q(z)\, d(H_n - H_0)(z) \right| &= \sup_{\mathcal{H}} \left| \int_{[0,\infty)} \frac{h(z)}{h(z)+h_0(z)}\, d(H_n - H_0)(z) \right| \\
\leq \ & \sup_{\mathcal{H}} \left| \int_{[0,\infty)} \frac{h(z)}{h(z)+h_0(z)} \mathbf{1}_{h_0(z)>\varepsilon}(z)\, d(H_n - H_0)(z) \right| \\
& + \sup_{\mathcal{H}} \left| \int_{[0,\infty)} \frac{h(z)}{h(z)+h_0(z)} \mathbf{1}_{h_0(z)\leq\varepsilon}(z)\, d(H_n - H_0)(z) \right| \\
\leq \ & \sup_{\mathcal{H}} \left| \int_{[0,\infty)} \frac{h(z)}{h(z)+h_0(z)} \mathbf{1}_{h_0(z)>\varepsilon}(z)\, d(H_n - H_0)(z) \right| \\
& + \sup_{\mathcal{H}} \int_{[0,\infty)} \mathbf{1}_{h_0(z)\leq\varepsilon}(z)\, d(H_n + H_0)(z) \\
\leq \ & \sup_{\mathcal{H}} \left| \int_{[0,\infty)} \frac{h(z)}{h(z)+h_0(z)} \mathbf{1}_{h_0(z)>\varepsilon}(z)\, d(H_n - H_0)(z) \right| \\
& + \int_{[0,\infty)} \mathbf{1}_{h_0(z)\leq\varepsilon}(z)\, d(H_n - H_0)(z) + 2 \int \mathbf{1}_{h_0(z)\leq\varepsilon}(z)\, dH_0(z). \quad (3.1.5)
\end{aligned}
$$

For any fixed $\varepsilon > 0$ another application of Theorem 3 of van der Vaart and Wellner (2000) using the continuous function $\varphi_2^\varepsilon : \mathbb{R}^2 \to \mathbb{R}$,

$$
\varphi_2^\varepsilon(x,y) = \begin{cases} 0 & \text{for } x < 0, y \in \mathbb{R}, \\ \frac{x}{x+y} & \text{for } x \geq 0, y > \varepsilon, \\ \frac{x}{x+\varepsilon} & \text{for } x \geq 0, y \leq \varepsilon, \end{cases}
$$

leads to $\varphi_2^\varepsilon(\mathcal{H}, \{h_0 \mathbf{1}_{\{h_0>\varepsilon\}}\})$ being H_0-Glivenko-Cantelli by setting $x = \mathcal{H}$ and $y = \{h_0\}$, i.e.

$$\mathbb{P}\left(\lim_{n\to\infty} \sup_{\mathcal{H}} \left| \int_{[0,\infty)} \frac{h(z)}{h(z)+h_0(z)} \mathbf{1}_{h_0(z)>\varepsilon}(z)\, d(H_n - H_0)(z) \right| = 0 \right) = 1. \quad (3.1.6)$$

The second term in (3.1.5) converges almost surely to zero by the strong law of large numbers since $\int \mathbf{1}_{h_0(z)\le\varepsilon}(z)\, dH_0(z) \le 1$, even for all $\varepsilon > 0$.

Now, denote the probability space supporting $Z_1, Z_2, \ldots$ by $(\Omega, \mathcal{S}, P)$. Let $\varepsilon_m = 1/m$ for $m \in \mathbb{N}$ and $\Omega^{\varepsilon_m} \in \mathcal{S}$ with $P(\Omega^{\varepsilon_m}) = 1$ for all $m \in \mathbb{N}$ such that the first and second term of (3.1.5) converge to zero for all $\omega \in \Omega^{\varepsilon_m}$. Furthermore, let $\Omega_0 = \bigcap_{m=1}^\infty \Omega^{\varepsilon_m}$ and note that $\mathbb{P}(\Omega_0) = 1$.

Fix $\omega \in \Omega_0$ and let $\eta > 0$. Choose $\varepsilon > 0$ such that the third term in (3.1.5) is smaller that $\eta/3$ which can be done by monotone convergence since $\mathbf{1}_{h_0(z)\le\varepsilon}(z)$ decreases for $\varepsilon \downarrow 0$. Then, by taking n sufficiently large, we have with probability equal to one,

$$\sup_{\mathcal{H}} \left| \int \frac{h(z)}{h(z)+h_0(z)}\, d(H_n - H_0)(z) \right| < \eta,$$

meaning that $\mathcal{H}^q$ is indeed a Glivenko-Cantelli class. □

The previous Lemma implies L_p-consistency of $\hat{h}_n$ for $p \ge 1$.

Corollary 3.1.3 (L_p-Consistency of $\hat{h}_n$).
Let $p \ge 1$. Then, for $n \to \infty$,

$$\|\hat{h}_n - h_0\|_{L_p} \rightarrow 0 \ a.s.$$

Proof.
For $p = 1$ this is an immediate consequence of (3.1.1) and Lemma 3.1.2.
For $p > 1$ note that $|a-b|^p \le |a-b|$ holds for $a, b \in [0,1]$. Then L_1-consistency implies

$$\left[\int \left| \frac{\hat{h}_n(z)}{g(0)} - \frac{h_0(z)}{g(0)} \right|^p dz \right]^{1/p} \rightarrow 0, \ a.s.$$

yielding L_p-consistency. □

We proceed by inferring in Theorem 3.1.5 (below) a consistency results for $\hat{F}_n$ from Lemma 3.1.2. In its proof the resolvent ϱ as introduced in Section 1.2 is used. We therefore state some properties of ϱ first.

LEMMA 3.1.4 (PROPERTIES OF ϱ).
Let g be a decreasing density with representation (1.2.3). *Then the resolvent ϱ defined as in Definition 1.2.1 is an increasing function with $\varrho(0) = g(0)^{-1}$ satisfying $\lim_{x\to\infty} \varrho(x)/x = 1$. Furthermore, due to* (1.2.3) *the function ϱ can be expressed as*

$$\varrho(x) \quad = \quad \frac{1}{g(0)} + \int_0^x \pi(w)\,dw,\ x \geq 0, \tag{3.1.7}$$

with a function $\pi : (0,\infty) \to [0,\infty)$ that is Lipschitz continuous on bounded intervals if g' is Lipschitz continuous. Moreover, ϱ is differentiable on $(0,\infty)$ with at most one exception at S_g.

PROOF.
See Jongbloed and van der Meulen (2008, Lemma 2.4) and Jongbloed (1995) for the last statement. □

THEOREM 3.1.5 (POINTWISE CONSISTENCY OF $\hat{F}_n$).
Let $x_0 > 0$ such that F_0 is continuous at x_0. Then the sequence $\hat{F}_n(x_0)$ converges almost surely to $F_0(x_0)$, i.e

$$|\ \hat{F}_n(x_0) - F_0(x_0)\ | \quad \to \quad 0 \text{ a.s. for } n \to \infty.$$

If F_0 is continuous, pointwise consistency of $\hat{F}_n$ can be strengthened to uniform consistency on $[0,\infty)$.

PROOF.
We first show pointwise consistency. Define, for $x \geq 0$,

$$T_F(x) \quad := \quad \int_0^x \varrho(x-y) h_F(y)\,dy \ = \ \int_0^x F(y)\,dy$$

where ϱ is the resolvent (see Definition 1.2.1). Note that T_F is bounded on bounded intervals, increasing and convex for all $F \in \mathcal{F}_{[0,\infty)}$. Using the notations $\hat{T}_n$ and T_0 for $T_{\hat{F}_n}$ and T_{F_0}, respectively, we get due to the L_1-consistency of $\hat{h}_n$ (Corollary 3.1.3) and monotonicity of ϱ (Lemma 3.1.4) for any $M > 0$ and $x \leq M$

$$\begin{aligned} |\ \hat{T}_n(x) - T_0(x)\ | &= \left| \int_0^x \varrho(x-y)\hat{h}_n(y)\,dy - \int_0^x \varrho(x-y) h_0(y)\,dy \right| \\ &\leq \ \varrho(x) \int_0^x |\ \hat{h}_n(y) - h_0(y)\ |\ dy \ \leq \ \varrho(M) \|\hat{h}_n - h_0\|_{L_1}. \end{aligned}$$

Since this upper bound in the last display is independent of x it implies for any open interval $I \subset [\,0, M\,]$

$$\sup_{x\in I} |\,\hat{T}_n(x) - T_0(x)\,| \quad \to \quad 0 \ a.s.$$

Now fix $x_0 > 0$ such that F_0 is continuous at x_0 and $M > x_0$. By the Lemma after Theorem 7.2.1 in Robertson et al. (1988) we now conclude that with probability one

$$F_0(x_0-) \;\le\; \liminf_{n\to\infty} \hat{F}_n(x_0-) \;\le\; \limsup_{n\to\infty} \hat{F}_n(x_0) \;\le\; F_0(x_0)$$

which can be restated as $\mathbb{P}\left(\lim_{n\to\infty} \hat{F}_n(x_0) = F_0(x_0)\right) = 1$ by the continuity of F_0 at x_0.

Assume that F_0 is continuous. Then one can define for any fixed $\eta > 0$ a finite grid on $[\,0,\infty)$, i.e. $0 = x_0 < x_1 < \ldots < x_k = \infty$ with $F_0(x_j) - F_0(x_{j-1}) < \eta$ for $j = 1,\ldots,k$. Thus, $\hat{F}_n$ converges uniformly to F_0 almost surely on the finite set $\{x_0,\ldots,x_k\}$. For $x > 0$ there exist a $j_0 \in \{1,\ldots,k\}$ such that $x_{j_0-1} \le x < x_{j_0}$ with

$$\hat{F}_n(x) - F_0(x) \quad\le\quad \hat{F}_n(x_{j_0}) - F_0(x_{j_0-1}) \;\le\; \hat{F}_n(x_{j_0}) - F_0(x_{j_0}) + \eta$$

and

$$\hat{F}_n(x) - F_0(x) \quad\ge\quad \hat{F}_n(x_{j_0-1}) - F_0(x_{j_0}) \;\ge\; \hat{F}_n(x_{j_0-1}) - F_0(x_{j_0-1}) - \eta$$

due to the monotonicity of $\hat{F}_n$ leading to the last statement of the theorem. □

Theorem 3.1.5 also implies uniform consistency of $\hat{h}_n$:

COROLLARY 3.1.6 (UNIFORM CONSISTENCY OF $\hat{h}_n$).
Let F_0 be continuous on $[\,0,\infty)$. Then the estimator $\hat{h}_n$ satisfies

$$\sup_{[\,0,\infty)} \left|\hat{h}_n(z) - h_0(z)\right| \quad\to\quad 0 \ a.s. \ for\ n \to \infty.$$

PROOF.
Due to (2.1.2) direct computations yield

$$\begin{aligned}
&\sup_{[\,0,\infty)} \left|\hat{h}_n(z) - h_0(z)\right| \\
&\quad\le\quad g(0) \sup_{[\,0,\infty)} \left|\hat{F}_n(z) - F_0(z)\right| + \sup_{y\in[\,0,\infty)} \left|\hat{F}_n(y) - F_0(y)\right| \sup_{z\in[\,0,\infty)} \int_0^z -g'(z-y)\,dy \\
&\quad=\quad 2g(0) \sup_{[\,0,\infty)} \left|\hat{F}_n(z) - F_0(z)\right|
\end{aligned}$$

which converges to zero almost surely by Theorem 3.1.5. □

3.2 An Asymptotic Property of $\hat{S}_n$

For this section we restrict ourselves to the case where F_0 is continuous with $S_0 = F_0^{-1}(1) < \infty$. Under this assumption, we study the asymptotic behavior of the estimator $\hat{S}_n = \inf\{x \geq 0 : \hat{F}_n(x) = 1\}$ of S_0.

The most desirable goal would be to show that $\hat{S}_n$ converges to S_0 in probably. This could be done in two parts by the following argument. Let $\varepsilon > 0$. Then, for all n,

$$\mathbb{P}\left(|\hat{S}_n - S_0| > \varepsilon\right) = \mathbb{P}\left(\hat{S}_n - S_0 > \varepsilon\right) + \mathbb{P}\left(\hat{S}_n - S_0 < -\varepsilon\right). \tag{3.2.1}$$

By applying Theorem 1 of Aarts et al. (2007) it follows that $\mathbb{P}\left(\hat{S}_n - S_0 > \varepsilon\right) \to 0$ as $n \to \infty$. Due to the above Theorem 3.1.5 the conditions of that theorem are met in the present setup.

At present we do not know whether the right side of (3.2.1) tends to zero as $n \to \infty$ for general g. However, it is true for a specific choice of g, namely $g(y) = \mathbf{1}_{[0,1]}(y)$. For more general g we can prove a weaker result under some local assumptions on g near zero although simulations show that one might be able to drop this assumption.

Uniform Deconvolution

For the uniform deconvolution, i.e. $g(y) = \mathbf{1}_{[0,1]}(y)$, consistency of $\hat{S}_n$ is established in Lemma 2 in Aarts et al. (2007). There it is shown that $\hat{S}_n \leq Z_n - 1 + V_n$ where Z_n denotes the largest order statistic and where $0 \leq V_n \leq O_p(\log n/n)$.

For the first term at the right side of (3.2.1) we then have, for $n \to \infty$,

$$\mathbb{P}\left(\hat{S}_n - S_0 > \varepsilon\right) \leq \mathbb{P}\left(Z_n - 1 + V_n - S_0 > \varepsilon\right) \leq \mathbb{P}\left(V_n > \varepsilon\right) \to 0$$

because of the behavior of V_n.

The proof of the cited Lemma 2 relies on the fact that the uniform noise density g is constant on $[0,1]$ and cannot be adapted to cover more general choices of g.

Decreasing Kernels with Compact Support

Without loss of generality (due to rescaling) we assume g to have compact support $[0,1]$. Furthermore we only consider those densities g with are continuously differentiable in a right neighborhood of zero and for which $-g'(0) < g(0)^2$. Note that the densities $g_m(y) = m(1-y)^{m-1}\mathbf{1}_{[0,1]}(y)$ satisfy this condition for all $m \in \mathbb{N}$.

The characterization of $\hat{F}_n$ given in Theorem 2.2.1 yields: $C_n(x) < 1 \implies x \notin \mathcal{T}_n$. Hence for every $b \in (0,1)$ we have that $\sup_{(S_0+b,S_0+1)} \hat{C}_n(x) < 1$ implies $\hat{S}_n \leq S_0 + b$. Thus, in order to establish consistency of $\hat{S}_n$ we would need to show that for all $\varepsilon > 0$

$$\mathbb{P}\left(\sup_{(S_0+\varepsilon,S_0+1)} \hat{C}_n(x) < 1 \right) \quad \to \quad 1,\ n \to \infty. \tag{3.2.2}$$

Note that the figures of $\hat{C}_n$ in Section 2.2 support this last statement. Nonetheless we can only show a similar but weaker result than (3.2.2) by considering the event $\left\{\hat{S}_n \leq S_0 + 1 - \delta\right\}$ for a small fixed δ instead of $\left\{\hat{S}_n \leq S_0 + \varepsilon\right\}$. By doing so, we are able to say that with probability tending to one $\hat{S}_n$ stays away from $S_0 + 1$. For the purpose of proving the local limit result in Chapter 5, the following Lemma is indeed sufficient.

Lemma 3.2.1.
Let g be a decreasing density on $[0,1]$ as in (2.1.3), satisfying $-g'(0) < g(0)^2$ and assume that g' is continuous in a right neighborhood of zero. Then there exists a small $\delta > 0$ and an $\eta > 0$ such that

$$\sup_{[S_0+1-\delta,S_0+1)} \hat{C}_n(x) \quad < \quad 1 - \eta$$

with probability tending to one as $n \to \infty$.

Proof.
Choose a small $\kappa > 0$ such that $-g'(0)/g(0)^2 < (1-\kappa)(1-4\kappa)/(1+\kappa)$ which is possible due to the assumption $-g'(0) < g(0)^2$. Furthermore, fix $x_0 \in (S_0, S_0+1)$ such that $g(S_0+1-x_0) \geq (1-\kappa)g(0)$ and $-g'(u) \leq -g'(0)(1+\kappa)$ for all $u \in [0, S_0+1-x_0]$ where the latter condition is guaranteed by the continuity of g'. Also, fix $x_1 < S_0 + 1$ satisfying $x_1 > x_0 + (1-3\kappa)(S_0+1-x_0)/(1-2\kappa)$.

Let $x \in [x_1, S_0+1)$. Then

$$\begin{aligned}\hat{C}_n(x) &= \hat{C}_n(x_0) + \hat{C}_n(x) - \hat{C}_n(x_0) \\ &= \hat{C}_n(x_0) + \int_{[x,S_0+1]} \frac{g(z-x) - g(z-x_0)}{\hat{h}_n(z)}\, dH_n(z) \\ &\qquad\qquad - \int_{[x_0,x)} \frac{g(z-x_0)}{\hat{h}_n(z)}\, dH_n(z).\end{aligned} \tag{3.2.3}$$

We proceed by considering each term of (3.2.3) separately. First recall that $\hat{C}_n(x_0) \leq 1$.

The second term of (3.2.3) can be written as

$$\begin{aligned}
&\int_{[x,S_0+1]} \frac{g(z-x)-g(z-x_0)}{\hat{h}_n(z)}\, dH_n(z) \\
&\le \int_{[x,S_0+1]} -g'(\xi_z)(x-x_0)\frac{g(z-x)}{g(S_0+1-x_0)}\frac{1}{\hat{h}_n(z)}\, dH_n(z) \\
&\le -g'(0)(1+\kappa)(x-x_0)\frac{1}{(1-\kappa)g(0)}\hat{C}_n(x) \;\le\; \frac{(1+\kappa)(-g'(0))}{(1-\kappa)g(0)}(S_0+1-x_0)
\end{aligned}$$

where ξ_z is contained in $[\,0, S_0+1-x_0\,]$ for all z. The third term of (3.2.3) can be bounded with probability tending to one in the following way since $x \ge x_1$:

$$\begin{aligned}
&-\int_{[x_0,x)} \frac{g(z-x_0)}{\hat{h}_n(z)}\, dH_n(z) \;\le\; -\int_{[x_0,x_1)} \frac{g(z-x_0)}{\hat{h}_n(z)}\, dH_n(z) \\
&\le -(1-\kappa)g(0)\int_{[x_0,x_1)} \frac{1}{\hat{h}_n(z)}\, dH_n(z) \;\le\; -(1-2\kappa)g(0)(x_1-x_0) \\
&\le -(1-3\kappa)g(0)(S_0+1-x_0)
\end{aligned}$$

for sufficiently large n where the second last inequality is due to Lemma 3.3.1 and the last inequality is implied by the choice of x_1. Thus, for all $x \in [\,x_1, S_0+1)$ we get with probability tending to one

$$\begin{aligned}
\hat{C}_n(x) &\le 1 + \frac{(1+\kappa)(-g'(0))}{(1-\kappa)g(0)}(S_0+1-x_0) - (1-3\kappa)g(0)(S_0+1-x_0) \\
&\le 1 + (1-4\kappa)g(0)(S_0+1-x_0) - (1-3\kappa)g(0)(S_0+1-x_0) \;<\; 1-2\eta
\end{aligned}$$

for $\eta = g(0)(S_0+1-x_0)\kappa/2 > 0$ by the definition of κ. Note that the right hand side of the previous display is independent of x which now allows the conclusion that

$$\sup_{[\,S_0+1-\delta, S_0+1)} \hat{C}_n(x) \;<\; 1-\eta$$

with probability tending to one by choosing $\delta = S_0+1-x_1$. □

Note that the above proof only works for densities g that are not too steep at zero, i.e. $-g'(0) < g(0)^2$. The following example illustrates that the result of Lemma 3.2.1 might still hold for choices of g that do not satisfy this condition.

Example.
The density $g^*(y) = (-60y+7)\,\mathbf{1}_{[\,0,5/53\,]}(y) + \left(\frac{-71}{48}y + \frac{71}{48}\right)\mathbf{1}_{(5/53,1\,]}(y)$ (see Figure 3.1) is piecewise linear, continuous and decreasing. Moreover, $-g^{*\prime}(0) = 60 > 49 = g^*(0)^2$ and $g^*(1) = 0$, thus the condition of Lemma 3.2.1 is violated.

For F_0 being the uniform distribution on $[0,1]$ we show $\hat{C}_{3000}$ (using a sample of size 3000 from $h_0 = g * F_0$) in Figure 3.2 which might suggests that the above Lemma might also hold for g^*. To get a better insight, we computed 50 upper support points S_n for different choices of g using $n = 1500$ observations. The bottom line ($y = 1$) in Figure 3.3 shows the 50 upper support points for $g_2(y) = 2(1-y)\mathbf{1}_{[0,1]}(y)$ for which the condition of Lemma 3.2.1 is satisfied and where one clearly sees that $\hat{S}_{1500}$ stays away from $S_0 + 1 = 2$. The middle line ($y = 2$) in the same figure shows the same situation for $g_3(y) = 3(1-y)^2\mathbf{1}_{[0,1]}(y)$ whereas the upper line ($y = 3$) uses g^*. Although it is not convincing from this picture that $\hat{S}_{1500}$ stays away from $S_0 + 1$ for g^*, it seems to be happening for very large n. Comparing $\hat{S}_{1500}, \hat{S}_{3000}$ and $\hat{S}_{10000}$ using only g^* as done from bottom to top in Figure 3.4 one can see that especially for 10000 observations $\hat{S}_{10000}$ concentrates around the true upper support point $S_0 = 1$.

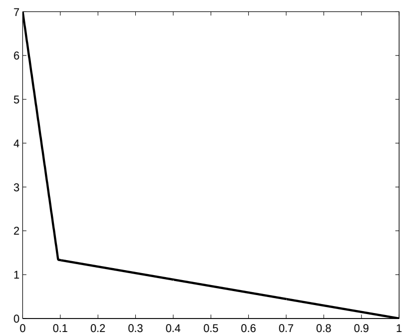

Figure 3.1: The density g^*.

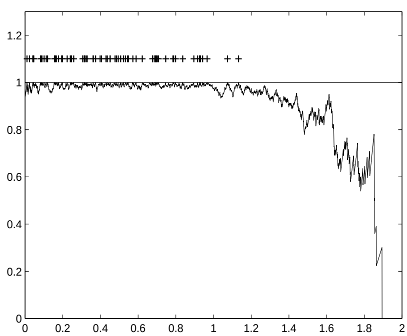

Figure 3.2: $\hat{C}_{3000}$ for g^*.

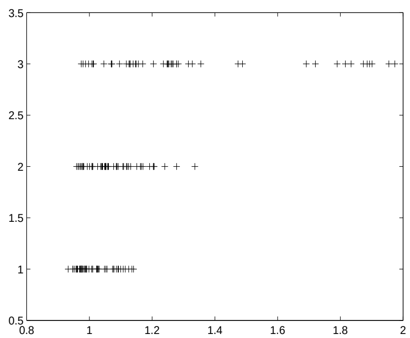

Figure 3.3: $\hat{S}_{1500}$ for different g.

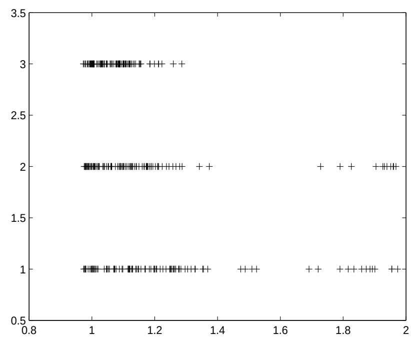

Figure 3.4: $\hat{S}_n$ for g^*.

3.3 Technical Lemma

Lemma 3.3.1.
Let $S_0 < a < b < S_0 + S_g < \infty$ such that $h_0(z) \geq c > 0$ for all $z \in [\,a, b\,]$ and some constant c. Then, for $n \to \infty$,

$$\int_{[\,a,b)} \frac{1}{\hat{h}_n(z)}\, dH_n(z) \quad \to \quad b - a \text{ a.s.}$$

Proof.
We write

$$\begin{aligned}\left|\int_{[\,a,b)} \frac{1}{\hat{h}_n(z)}\, dH_n(z) - b + a\right| &= \left|\int_{[\,a,b)} \frac{1}{\hat{h}_n(z)}\, dH_n(z) - \int_{[\,a,b)} \frac{1}{h_0(z)}\, dH_0(z)\right| \\ &\leq \left|\int_{[\,a,b)} \left(\frac{1}{\hat{h}_n(z)} - \frac{1}{h_0(z)}\right) dH_n(z)\right| + \left|\int_{[\,a,b)} \frac{1}{h_0(z)}\, d(H_n - H_0)(z)\right| \\ &\leq \frac{\|\hat{h}_n - h_0\|_{L_\infty}}{c} \int_{[\,a,S_0+S_g)} \frac{1}{\hat{h}_n(z)}\, dH_n(z) + \left|\int_{[\,a,b)} \frac{1}{h_0(z)}\, d(H_n - H_0)(z)\right| \qquad (3.3.1)\end{aligned}$$

with $\|\hat{h}_n - h_0\|_{L_\infty} = \sup_{z \in [\,0,\infty)} |\hat{h}_n(z) - h_0(z)|$.
First note that due to Theorem 2.2.1

$$\begin{aligned}\int_{[\,a,S_0+S_g)} \frac{1}{\hat{h}_n(z)}\, dH_n(z) &\leq \frac{1}{g(S_0+S_g-a)} \int_{[\,a,S_0+S_g)} \frac{g(z-a)}{\hat{h}_n(z)}\, dH_n(z) \\ &\leq \frac{1}{g(S_0+S_g-a)}.\end{aligned}$$

Combining this with uniform consistency of $\hat{h}_n$ (see Corollary 3.1.6), it follows that the first term in (3.3.1) converges to zero almost surely. The second term in (3.3.1) also tends to zero by the strong law of large numbers since $\int_{[\,a,b)} h_0^{-1}(z)\, dH_0(z) \leq S_0 + S_g < \infty$ which finishes the proof. □